28. Osteoporosis in Postmenopausal Women: Diagnosis and Monitoring

Evidence Report/Technology Assessment

Number 28

Prepared for:
Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
2101 East Jefferson Street
Rockville, MD 20852
http://www.ahrq.gov

Contract No. 290-97-0018

Prepared by:
Oregon Health & Science University Evidence-based Practice Center
Portland, Oregon
Heidi D. Nelson, M.D., M.P.H.
Principal Investigator

Cynthia D. Morris, Ph.D., M.P.H.
Dale F. Kraemer, Ph.D.
Susan Mahon, M.P.H.
Nancy Carney, Ph.D.
Peggy M. Nygren, M.A.
Mark Helfand, M.D., M.P.H., EPC Director

AHRQ Publication No. 01-E032

November 2001

Suggested Citation:
Nelson HD, Morris CD, Kraemer DF, et al. Osteoporosis in postmenopausal women: diagnosis

and monitoring. Evidence Report/Technology Assessment No. 28 (Prepared by the Oregon Health & Science University Evidence-based Practice Center under Contract No. 290-97-0018). AHRQ Publication No. 01-E032. Rockville, MD: Agency for Healthcare Research and Quality. January 2001. ↑ TOP

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-Based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To bring the broadest range of experts into the development of evidence reports and health technology assessments, AHRQ encourages the EPCs to form partnerships and enter into collaborations with other medical and research organizations. The EPCs work with these partner organizations to ensure that the evidence reports and technology assessments they produce will become building blocks for health care quality improvement projects throughout the Nation. The reports undergo peer review prior to their release.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality.

We welcome written comments on this evidence report. They may be sent to: Acting Director, Center for Practice and Technoloy Assessment, Agency for Healthcare Research and Quality, 6010 Executive Blvd., Suite 300, Rockville, MD 20852.

John M. Eisenberg, M.D.	Director, Center for Practice and
Director, Agency for Healthcare	Technology Assessment
Research and Quality	

The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S.

Robert Graham, M.D. ↟ TOP

Acknowledgments

Structured Abstract

Objectives.

This report examines the evidence on the effectiveness of various strategies for diagnosing and monitoring postmenopausal women with osteoporosis. Specifically, it addresses: (1) the role of risk factors in identifying high-risk women and guiding their initial treatment, (2) the advantages and disadvantages of various techniques for bone measurement in predicting risk of hip or spine fracture, (3) the effectiveness of bone measurement tests for monitoring response to treatment and for guiding treatment change, (4) the role of markers of bone turnover in diagnosis and treatment management, (5) the evaluation of patients with osteoporosis for secondary causes, and (6) the costs and cost-effectiveness of various diagnostic strategies for osteoporosis. ↟ TOP

Search Strategy.

The authors conducted a MEDLINE search (covering the years 1966 to 2000), supplemented by searches of HealthSTAR (covering 1975 to 2000) of papers published in English, reviewed reference lists of review articles, and sought guidance from local and national experts. ↟ TOP

Selection of Studies.

The authors included abstracts relevant to one or more topic areas that had original data about postmenopausal women and osteoporosis. Two reviewers read each abstract to determine its eligibility. Articles were excluded if they did not provide sufficient information to determine the methods for selecting subjects and for analyzing data. For some topics, additional eligibility criteria were applied. For all topics combined, the authors retrieved 10,174 citations. After reviewing these citations for possible relevance, 530 articles about risk factors, 123 about bone measurement testing, 23 about monitoring, 277 about biochemical markers, and 53 about costs were selected for further review. An additional 242 studies were retrieved after reviewing the reference lists of studies and/or by suggestion of others. The search yielded no papers with data for the secondary causes topic. ⬆ TOP

Data Collection and Analysis.

From full-text published studies of fracture or bone density prediction or bone measurement methods, the authors extracted selected information about the patient population, interventions, clinical endpoints, study design, and study quality, and used this information to construct evidence tables. Additional reviews assessed the internal validity of studies of risk factors and the diagnostic performance of bone measurement tests and biochemical markers, summarized recommendations for testing for secondary causes of osteoporosis, reviewed studies about cost and cost-effectiveness, and compared diagnostic strategies. ⬆ TOP

Main Results.

Epidemiologic studies report clinical risk factors for osteoporosis and fractures, but few studies evaluate how to use them to identify individual women at risk for fracture, and no studies provide evidence that treatment decisions based on clinical risk factors lead to better or worse fracture outcomes than those based on bone measurement tests. Because of differences between bone measurement techniques, and because individuals have different rates of bone loss at different sites, no one test can exclude osteoporosis at the most important fracture sites -- hip, spine, and wrist. Dual-energy X-ray absorptiometry (DXA) of the femoral neck is the best validated test to predict hip fracture. Other techniques predict hip fracture less accurately or have not been evaluated in prospective studies.

Recent results from clinical trials raise questions about the value of repeated, annual densitometry tests for patients on therapy to prevent osteoporosis or bone loss; moreover, there is no evidence from clinical trials that adjusting therapy based on serial densitometry at any interval improve outcomes. Markers of bone turnover correlate poorly with bone measurement tests and are not good predictors of fractures.

Cost and cost-effectiveness studies, which are based solely on economic models, suggest targeting treatment to women with the lowest bone density and including a risk factor score or less expensive (and more widely available) technology to determine which women should receive hip DXA. The authors' supplementary analysis on cost-effectiveness favors a sequential strategy of quantitative ultrasound at the heel followed by densitometry of those identified by ultrasound as high risk over densitometry alone. In high-risk populations, ultrasound alone may

also be cost-effective. ↑ TOP

Conclusions.

Application of the results from epidemiologic studies to diagnosis and monitoring strategies for individual patients in the clinical setting is currently based on extrapolation from models or, for most questions addressed in this review, is simply lacking. To be more useful for clinicians and patients, future research should focus on the application of these data to the clinical setting.
↑ TOP

Summary

Overview

At an international consensus development conference, osteoporosis was defined as "a systemic skeletal disease characterized by low bone mass and microarchitectural deterioration of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk." In 1994, a World Health Organization (WHO) working group proposed that, in epidemiologic studies, osteoporosis could be determined when bone density at the hip, spine, or forearm is 2.5 standard deviations or more below the mean for healthy, young, adult women (a value defined as the T-score), or when a history of a fracture is present in the absence of trauma. The group also proposed that osteopenia be diagnosed when the bone density was 1.0 to 2.5 standard deviations below the mean for young, healthy women.

According to the National Health and Nutrition Examination Survey (NHANES III), an estimated 14 million American women over age 50 years are affected by low bone density at the hip, and 5 million more have bone density that measures -2.5 standard deviations or more below the mean at the hip. The prevalence of osteoporosis in Mexican-American women is similar to that in white women, while rates in black women are approximately half that of the first two groups. The prevalence of osteoporosis increases with age for all sites, and -- by the WHO definition -- up to 70 percent of women over age 80 years have osteoporosis. Furthermore, age is an important factor in the relationship between bone density and the absolute risk of fracture. An increase in age of 13 years increases the risk of hip fracture by the same amount as a decrease in bone density of one standard deviation. Older women have a much higher fracture rate than younger women who have the same bone density, because of increasing risk from other factors, such as a change in bone quality and the tendency to fall.

Women with osteoporosis are more likely to experience fractures. Demographic trends for hip fracture parallel those for osteoporosis. Hip-fracture incidence in white women rises from 50 per 100,000 at age 50 years to 237 per 100,000 at age 65 years. White women are generally two to three times more likely than nonwhite women to suffer a hip fracture. Hip fractures are associated with high rates of mortality and loss of independence. Wrist fracture incidence tends to increase at earlier ages than does that of hip fractures.

Vertebral fractures have also been associated with significant morbidity. Sixteen percent of postmenopausal women have osteoporosis of the lumbar spine; furthermore, five percent of 50-

year-old white women and 25 percent of 80-year-old women have had at least one vertebral fracture. Vertebral fractures can cause severe pain and are associated with more than five million days of restricted activity in those age 45 years or older.

The disease burden of osteoporosis extends beyond consequences of low bone density and fractures. For example, the act of screening, diagnosis, and subsequent treatment can also affect the quality of life. Fear of fracture itself can reduce the quality of life in women who have been diagnosed as having osteoporosis.

In 1995, the total direct medical expenditure in the United States for the treatment of osteoporotic fractures in adults older than 45 years was estimated at $13.8 billion. The majority of this total ($8.6 billion) was spent for inpatient care. Hip fracture alone accounted for $8.7 billion (63 percent) of osteoporosis-related costs, while fractures at sites other than the hip accounted for approximately 37 percent of the total expenditure (about $5.1 billion). In addition to these costs is the cost of lost productivity for women with fractures, or for their family or other caregivers. As the median age of the U.S. population increases, the costs associated with osteoporotic fractures are also likely to increase. ↑ TOP

Reporting the Evidence

This evidence report describes the effectiveness of various strategies for diagnosing and monitoring postmenopausal women with osteoporosis, as represented by six topic areas:

1. *Risk Factors*. What is the role of clinical risk factors, in conjunction with bone measurement tests, in identifying high-risk women and guiding initial treatment decisions?
2. *Bone Measurement Tests*. What are the advantages and disadvantages of various bone measurement tests at different anatomic sites for identifying women at high risk of fracture?
3. *Monitoring*. Are bone measurement tests effective for monitoring response to treatment and for guiding decisions about changes in management?
4. *Biochemical Markers*. What is the role of markers of bone turnover for identifying women at risk of bone loss, guiding initial treatment decisions, or monitoring response to therapy?
5. *Evaluation for Secondary Causes*. What diagnostic or laboratory tests are appropriate for evaluating patients with osteoporosis for secondary causes?
6. *Cost*. Assuming consistent treatment approaches, what are the costs and cost-effectiveness of various diagnostic strategies for identifying women with osteoporosis?

These topic areas do not include the effectiveness of dietary, lifestyle, hormonal, and medical interventions for primary prevention or treatment of osteoporosis. This report is confined to diagnostic and monitoring strategies as they apply to individual women, and does not include issues regarding mass screening in the general population. Also, while most of the literature addressing these topic areas is aimed at an audience of clinical researchers who specialize in osteoporosis, we have attempted to assess the research findings from the perspectives of clinicians and patients. However, it is not the purpose of this report to propose practice

recommendations. ⬆ TOP

Methodology

With input from local and national experts, we developed an analytic framework and key questions in each of the six topic areas. Relevant studies were identified from multiple searches of MEDLINE (for the years 1966 to 2000) and HealthSTAR (for the years 1975 to 2000), from the reference lists of systematic review articles, and from national experts. All searches were limited to publications in the English language. The authors excluded articles if they did not provide sufficient information to determine the methods for selecting subjects and for analyzing data.

Out of 10,174 citations retrieved for all topics combined, the authors selected as possibly relevant 530 articles about risk factors, 123 about bone measurement tests, 23 about measurement tests for bone monitoring, 277 about biochemical markers, and 53 about costs. An additional 242 studies were retrieved after reviewing reference lists of studies and by suggestion of the expert panel or leading researchers in the field. The search yielded no papers with data for the secondary causes topic, but this topic was addressed by reviewing published guidelines and by a supplemental analysis of physician practice patterns for the evaluation of secondary causes of osteoporosis. For articles that were included, the authors applied criteria proposed by the U.S. Preventive Services Task Force to rate the quality of individual studies. A second supplemental analysis investigated the cost-effectiveness of various diagnostic strategies. ⬆ TOP

Findings

Risk Factors

After a review of the relevant articles, the findings about the risk factors that predict bone density, bone loss, and fractures are as follows:

- Factors that are consistently associated with increased risks of low bone density and fractures in postmenopausal women include increasing age, white race, low weight or weight loss, nonuse of estrogen replacement, history of previous fracture, family history of fracture, history of falls, and low scores on one or more measures of physical activity or function.
- Other factors are less consistent predictors across studies, but also have statistically significant associations with low bone density and fractures. These include smoking, alcohol use, caffeine use, low calcium and vitamin D intake, and use of certain drugs.
- Predictors of low bone density are similar to those for fracture, except for those factors related to physical function and falls.
- Some clinical risk factors, especially those related to physical function and falls, are as powerful as bone density in the prediction of hip fracture.
- Women with multiple risk factors and low bone density have an especially high risk of hip fracture.
- Most of the strongest risk factors are consistently related to outcomes, regardless of the racial and ethnic population, although there are few studies of nonwhite women.

A second set of findings about risk factors concerned the accuracy of methodologies (also called instruments or tools) for assessing risk factors in identifying women at risk of fracture. These included:

- In contrast to the extensive research about determining clinical risk factors for osteoporosis and fractures, there are fewer studies available that evaluate how to use these risk factors to identify individual women at risk for fracture.
- Methologies designed to assess risk of low bone density or fractures generally have low to moderate sensitivity and specificity. Those that perform well when originally tested either have performed less well in other populations or have not yet been widely tested. Some methodologies -- especially those developed in large community populations and containing variables known to be strong predictors -- may ultimately be applicable to the clinical setting once they are tested there.

A third set of findings, exploring whether risk factors are useful in treatment decisions, included:

- The authors did not identify any studies that examined whether treatment decisions based on clinical risk factors lead to better or worse health outcomes than those based on bone measurement tests or a combination of bone tests and risk factors.

Bone Measurement Tests

Regarding bone measurement tests, the major findings related to the capacity of different bone measurement tests at different sites to predict fractures included:

- Among different bone measurement tests at various sites, bone density measured at the femoral neck by dual energy x-ray absorptiometry (DXA) is the best predictor of hip fracture and is comparable to forearm measurements for predicting fracture at other sites.
- In recent prospective studies, quantitative ultrasound (QUS) measured at the heel predicted hip fracture and all nonspine fractures as well or nearly as well as DXA measured at the femoral neck. For each of these tests, a result in the osteoporotic range is independently associated with an increased short-term probability of hip fracture. Individuals who have low scores by one of these tests, but not the other, have a higher risk of fracture than those who have higher scores by both tests, and a lower risk of fracture than those whose results on both tests are low.
- While other peripheral measures may approach QUS in predicting hip fracture, there are no recent prospective studies that directly compare prediction of hip fracture of these tests with DXA of the hip. Radiographic absorptiometry (RA) or quantitative microdensitometry (QMD) of the hand can predict the risk of nonspine fractures in general, many of which are in the forearm, but there are no recent data about the ability of hand measurement to predict hip fracture.
- Correlations between different bone measurement tests are generally too low to be accepted as evidence that one test will identify patients at similar risk to those identified by another test.

Major findings on identification of the factors related to bone testing that influence diagnosis included:

- The likelihood of being diagnosed with osteoporosis varies greatly, depending on the site and type of the bone measurement test, on the brand of densitometer, and on the relevance of the reference range to the local population.
- The likelihood of being diagnosed as having osteoporosis also depends on the number of sites tested. Testing in the forearm, hip, spine, or heel will generally identify different groups. A physician cannot say, based only on one of these tests, that the patient "does not have osteoporosis."
- The results of bone measurement tests are often inaccurately reported to patients.

Major findings related to how bone measurement test results affect patients' and physicians' decisions and actions included:

- One randomized trial suggests that women who undergo densitometry are more likely to start hormone replacement therapy than women who do not.
- In a randomized trial and a large, uncontrolled case series, women who had densitometry and were told they had osteoporosis were more likely to start or continue hormone replacement therapy than women who were told they had normal bone density.
- In one randomized trial, physicians found densitometry reports confusing and were not confident that their interpretations of T-scores were correct.

Monitoring

The major findings regarding whether bone measurement tests are effective for monitoring response to therapy and for guiding decisions about changes in management included:

- The weight of the evidence is currently against repeating bone density tests within the first year of treatment. There is insufficient evidence to determine whether repeating bone density tests 2 years after starting therapy is useful.
- There are also no studies about the effect of either monitoring responses to therapy using densitometry or the choice of test on the outcome of therapy.

Biochemical Markers

Findings regarding whether biochemical markers can be used instead of bone measurement tests in identifying women at risk for osteoporosis included:

- No single marker or cluster of markers accurately predicted the results of densitometry in individuals. Densitometry measures current bone status, whereas markers measure the process of bone turnover.

Major findings as to how well biochemical markers predict fracture included:

- No marker was associated with increased fracture risk consistently across all studies. One study provides evidence that using markers in conjunction with densitometry may increase predictability, but this result has not been otherwise confirmed.

In addition, major findings as to whether markers can help select patients for treatment included:

- Studies correlating marker results and bone loss indicated no clear trend. Furthermore, sensitivity and specificity of markers were too low to be useful for the purpose of selecting patients for treatment.
- Some studies found better test accuracy when a combination of two or more markers and/or other risk factors was used to predict bone loss.

The report also investigated what is known about the adverse effects of using markers to identify women at risk for osteoporosis. Major findings for this issue included:

- The primary adverse effect of biochemical markers is the potential for false-positive and false-negative results. Rates of false-positive and false-negative test results vary widely among the biochemical markers. (If markers are used to select women for treatment, a false-positive test could lead to the initiation of unnecessary treatment, while a false-negative result could lead to lack of needed treatment.)

Major findings regarding whether markers can predict a patient's response to therapy included:

- There is a small correlation between response to therapy as measured by densitometry and marker results, but no marker is accurate enough to reliably identify those individuals who will fail to respond to treatment.

Evaluation for Secondary Causes of Osteoporosis

Major findings from a literature review, assessment of published guidelines, and a secondary analysis regarding which diagnostic or laboratory tests are appropriate for evaluating patients with osteoporosis for secondary causes included:

- There is no evidence from controlled trials on which to base recommendations for a strategy of testing to determine secondary causes of osteoporosis.
- Some guidelines and experts support extensive testing to rule out major concomitant disease, while others suggest a limited testing strategy based on findings in the history and physical examination. Because the diagnosis of primary osteoporosis is often seen as a diagnosis of exclusion, the pattern of diagnostic testing may continue to be costly until the diagnostic yield is fully demonstrated.
- Assumptions about the probability of a secondary disease or disorder to explain the occurrence of osteoporosis vary by practice type and specialty.
- Thyroid stimulating hormone measurement, chemistry battery, and complete blood count were the most frequently ordered tests to rule out secondary causes of osteoporosis cited by respondents in our supplemental analysis. These were also the most frequently

recommended tests in our review of expert guidelines.

TOP
Costs

The report's findings regarding the costs and cost-effectiveness of diagnostic strategies for identifying women with osteoporosis included:

- Published economic assessments suggested that diagnosis and treatment of women at risk for osteoporosis would be more cost-effective by targeting treatment to those with the lowest bone measurement results. Inclusion of another assessment, such as a risk profile or additional bone measurement test, prior to DXA may improve the cost-effectiveness of diagnosis.
- Using data from two large studies, the authors conducted cost-effectiveness analyses to estimate cost per hip fracture prevented. These analyses suggested that a sequential diagnostic approach may be more cost-effective than DXA alone. The sequential approach that the authors considered was QUS of the heel followed by DXA of the femoral neck only for those with low values of QUS/BUA. A range of QUS/BUA measures was used because there are no established cut points that separate high risk from normal risk. The authors used QUS as an example of a less expensive and more widely available diagnostic approach than DXA.
- Diagnosis with DXA of the femoral neck alone prevented more fractures, in most cases, but at a higher cost per hip fracture prevented compared with the sequential approach.
- All current treatment trials have inclusion criteria based on diagnosis with DXA. In sensitivity analyses, if treatment efficacy following diagnosis with QUS (using QUS/BUA cut points between 65 and 75 dB/mHz) were 5%-15% less than that following diagnosis using DXA, diagnosis with QUS alone would have higher costs to prevent fewer hip fractures than other diagnostic options.

TOP
Future Research

Much of the evidence for the diagnosis and monitoring strategies for osteoporosis comes from epidemiologic studies. To be more useful for clinicians and patients, future research should focus on the application of these data to the clinical setting and include a wider diversity of patient populations. Tools for assessing risk factors should be tested in prospective studies to determine if their use can correctly stratify women by risk factors, influence treatment decisions, and ultimately reduce fracture outcomes.

Clinical trials should be conducted to determine if identifying and reducing modifiable risk factors influences fracture outcomes. Addressing some of these modifiable risk factors -- such as by supplementation with calcium and vitamin D -- has already demonstrated effects on fracture risk after intervention. Examples of additional interventions to test include smoking cessation, correcting visual loss, and improving physical function. Randomized, controlled trials of treatments for osteoporosis should be done to test the hypothesis that overall fracture risk, rather than bone measurement results alone, determines the likelihood that a patient will benefit from

therapy. Selection of patients for trials should focus on groups of patients who are at high risk of fracture based on clinical risk factors and who have relatively normal bone measurement results. Trials should also address whether patients who are identified to have bone loss demonstrated by different techniques at different sites demonstrate a similar benefit to those identified solely by hip DXA measurements.

Additional research is needed to examine the quality of information provided to patients who undergo various bone measurement tests, as well as to identify other patient education and information needs.

Future research should examine the clinical utility of the WHO working group's criterion for diagnosing osteoporosis, specifically:

- To determine whether the diagnosis of osteoporosis provides benefits to patients above that provided by predicting their risk of fractures;
- To assess the added value of obtaining bone measurement tests at more than one site, and the usefulness of using T-scores for different sites; and
- To define the prognosis and treated course in patients who are diagnosed to have osteoporosis by a wrist or spine measure, but who are not diagnosed to have osteoporosis by a hip measure.

Studies that use bone measurement tests for monitoring response to therapy could compare fracture outcomes in a group of patients who had tailored therapy based on test results versus a group in whom changes in therapy, if any, were guided by the history alone. A study could also record how often test results in patients on therapy led to a change in therapy or improved compliance, to establish the mechanism by which monitoring leads to improved outcomes. By comparing patients, such a study could establish that monitoring changes in test results can reliably predict fracture risk in individual patients by distinguishing an inadequate response to therapy from an adequate response or poor adherence, and that monitoring changes in therapy made because of an inadequate test response can reduce the rate of fractures.

Prospective studies of biochemical markers should define, apply, and evaluate criteria for using marker results in clinical decision making.

Determining the utility of screening for secondary disorders by use of common laboratory tests requires studies of the frequency of abnormal baseline laboratory test results in large cohorts or in treatment trials of osteoporosis. Clinical follow-up of these subjects would provide data on bone measurement test results or fracture outcomes.

Studies to formally account for the adverse quality-of-life impact of treatment and treatment side effects are needed to more accurately determine the balance of risks and benefits of the therapy options for patients.

Three areas for additional cost-effectiveness research include:

- Identifying a scientifically appropriate cut point for QUS/BUA that can be used in either

a sequential diagnostic approach or for diagnosis with QUS alone.
- Performing additional cost-effectiveness analyses using data from other large, population-based cohorts with various cut points.
- Conducting a more detailed, societal-perspective, cost-effectiveness analysis that would address some of the deficiencies in the authors' analysis.

If these findings can be demonstrated in one or two other populations, and in a more complete economic evaluation, a randomized, controlled trial of diagnostic approaches would be a useful next step. Alternatively, an observational-design study with randomization of groups of patients to diagnostic or monitoring procedures could be conducted. ↑ TOP

[Copyright and Disclaimer](#)